


Evaluation of fibrosis and inflammation in diffuse liver diseases using intravoxel incoherent motion diffusion-weighted MR imaging

Manuela França ¹, Luis Martí-Bonmatí,² Ángel Alberich-Bayarri,² Pedro Oliveira,^{3,4} Susana Guimaraes,⁵ João Oliveira,⁶ João Amorim,⁶ Javier Sanchez Gonzalez,⁷ José Ramón Vizcaíno,⁸ Helena Pessegueiro Miranda^{4,9}

¹Imaging Department, Centro Hospitalar do Porto, Largo Prof Abel Salazar, 4099-001 Porto, Portugal

²Radiology Department, Hospital Universitario y Politécnico La Fe and Biomedical Imaging Research Group (GIBI230), Valencia, Spain

³Population Studies Department, Institute of Biomedical Sciences Abel Salazar (ICBAS), University of Porto, Porto, Portugal

⁴Epidemiology Research Unit (EPI Unit), Institute of Public Health of the University of Porto, Porto, Portugal

⁵Pathology Department, Centro Hospitalar de S. João, Porto, Portugal

⁶Radiology Department, Centro Hospitalar do Porto, Porto, Portugal

⁷MR Clinical Science, Philips Healthcare, Madrid, Spain

⁸Pathology Department, Centro Hospitalar do Porto, Porto, Portugal

⁹Liver and Pancreas Transplantation Unit and Medicine Department, Centro Hospitalar do Porto, Porto, Portugal

Abstract

Purpose: The purpose of the study was to evaluate the role of intravoxel incoherent motion (IVIM) diffusion model for the assessment of liver fibrosis and inflammation in diffuse liver disorders, also considering the presence of liver steatosis and iron deposits.

Methods: Seventy-four patients were included, with liver biopsy and a 3 Tesla abdominal magnetic resonance imaging examination, with an IVIM diffusion-weighted sequence (single-shot spin-echo echo-planar sequence, with gradient reversal fat suppression; 6 b -values: 0, 50, 200, 400, 600, and 800 s/mm²). Histological evaluation comprised the Ishak modified scale, for grading inflammation and fibrosis, plus steatosis and iron loading classification. The liver apparent diffusion coefficient (ADC) and IVIM parameters (D , D^* , f) were calculated from the IVIM images. The relationship between IVIM parameters and histopathological scores were evaluated by ANOVA and Spearman correlation tests. A test–

retest experiment assessed reproducibility and repeatability in 10 healthy volunteers and 10 randomly selected patient studies.

Results: ADC and f values were lower with higher fibrosis stages ($p = 0.009$, $p = 0.006$, respectively) and also with higher necro-inflammatory activity grades ($p = 0.02$, $p = 0.017$, respectively). Considered together, only fibrosis presented a significant effect on ADC and f measurements ($p < 0.05$), whereas inflammation had no significant effect ($p > 0.05$). A mild correlation was found between ADC and f with fibrosis ($R_s = -0.32$ and $R_s = -0.38$; $p < 0.05$) and inflammation ($R_s = -0.31$ and $R_s = -0.32$, $p < 0.05$; respectively). The AUROC for ADC and f measurements with the different dichotomizations between fibrosis or inflammation grades were only fair (0.670 to 0.749, $p < 0.05$). Neither D nor D^* values were significantly different between liver fibrosis or inflammation grades. D measurements were significantly different across histologic grades of steatosis ($p < 0.001$) and iron overload ($p < 0.001$), whereas f measurements showed significant differences across histologic steatosis grades ($p = 0.005$). There was an excellent agreement between the different readers for ADC, f , and D .

Conclusions: Although fibrosis presented a significant effect on ADC and f , IVIM measurements are not accurate enough to stage liver fibrosis or necro-inflammatory activity in diffuse liver diseases. D values were influenced by steatosis and iron overload.

Key words: Magnetic resonance—Diffusion-weighted imaging—Intravoxel incoherent motion—Liver fibrosis—Liver steatosis—Iron overload

Diffuse liver diseases are related to a wide spectrum of etiologies, such as viral hepatitis infection, alcohol and drug consumption, fatty liver, auto-immune and metabolic diseases, leading to longstanding liver damage and inflammatory reaction, and finally triggering fibrosis with extracellular sinusoidal accumulation of collagens, proteoglycans and macromolecules. Patients with liver inflammation and fibrosis may remain asymptomatic, being important to promptly and reliably detect and grade both conditions to start treatment early enough [1]. Although liver biopsy is still considered the gold standard for the evaluation of liver fibrosis and inflammation, its use is declining, as the technique is invasive and might not be safe for the patient [2]. Furthermore, biopsy has accuracy limitations related to the large sampling bias and high inter- and intra-observer variability reports [3]. Consequently, there is a need to evaluate noninvasive biomarkers for diagnosing and staging the whole liver inflammatory injury and fibrosis status.

Diffusion-weighted (DW) magnetic resonance (MR) imaging has been proposed as a tool to assess liver inflammation and fibrosis [4–9]. Apparent diffusion coefficient (ADC), derived from the monoexponential fitting model, is heavily influenced by the b -value choice and it is influenced not only by the diffusion of water in the extracellular space but also by the capillary blood perfusion [10]. The bi-exponential intravoxel incoherent motion (IVIM) DW model [11, 12] allows separating the diffusion and perfusion effects by quantifying the pure diffusion coefficient (D), the pseudo-diffusion (D^*), and vascular fraction (f) microperfusion parameters. Lower ADC values in liver with fibrosis were associated to both liver compartment restrictions [13–16]. Other authors found lower ADC values in fibrotic livers being only dependent on perfusion parameters changes [5, 17, 18].

Liver inflammation is characterized by different grades of hepatocytes' swelling and necrosis, inflammatory cells migration through liver sinusoids and increased regional blood flow [19]. Inflammatory responses may also influence ADC measurements [6, 7, 9, 20–22] and both diffusion and perfusion IVIM parameters.

The presence of concomitant fat and iron deposits in the liver parenchyma might also affect water diffusion and the derived biomarkers [23–27]. Moreover, the DW

sequence, performed with an echo-planar imaging technique, is prone to susceptibility artifacts induced by iron deposits [23]. The possible interaction and confounding effects of liver inflammation, fibrosis, fat and iron deposits in the IVIM parameters have not been previously addressed in diffuse liver diseases.

Our objective was to evaluate the diagnostic accuracy of IVIM parameters for the detection and staging of both liver fibrosis and inflammation, considering the different clinical scenarios of diffuse liver disorders, and taking into consideration the effects of steatosis and iron overload.

Patients and methods

Study design and population

The Institutional Review Board approved this prospective, comparative, non-randomized single center study.

Between April 2013 and January 2014, all adult patients referred for liver biopsy were invited to participate in this study, except those with contra-indications to MR examination. Over these 10 months, 88 patients were recruited after obtaining written informed consent from them. Since severe iron overload lowers the liver signal intensity on MR images, patients with known hemochromatosis were excluded ($n = 8$). Six patients were afterwards excluded due to MR imaging movement artifacts ($n = 3$), malignancy within the biopsy sample ($n = 2$) and claustrophobia ($n = 1$). The final series included 74 patients (Fig. 1).

To check for MR reproducibility and repeatability, 10 healthy volunteers with no liver disease were also collected and evaluated with the same MR sequences.

Liver biopsy

One percutaneous biopsy sample per patient was obtained under ultrasound guidance, using 16–18G needles. Patients with focal liver lesions and suspected of having diffuse liver disease underwent a second targeted biopsy for the lesion. Although most biopsies were sampled from the right anterior segments, the Couinaud segmentation was registered to locate the regions of interest (ROIs) at the MR images. In cases of patients that underwent liver biopsy after the MR examination (10 patients), the liver biopsy was performed blinded to MR evaluation and the MR imaging analysis and ROI measurements were performed after biopsy.

Histological evaluation

Histological evaluation of biopsy samples was considered the reference standard for the prospective assessment and grading of liver inflammation and fibrosis, steatosis and iron deposits.

One pathologist (SG, 9 years' experience), blinded to the MR results, verified the quality of the specimens and

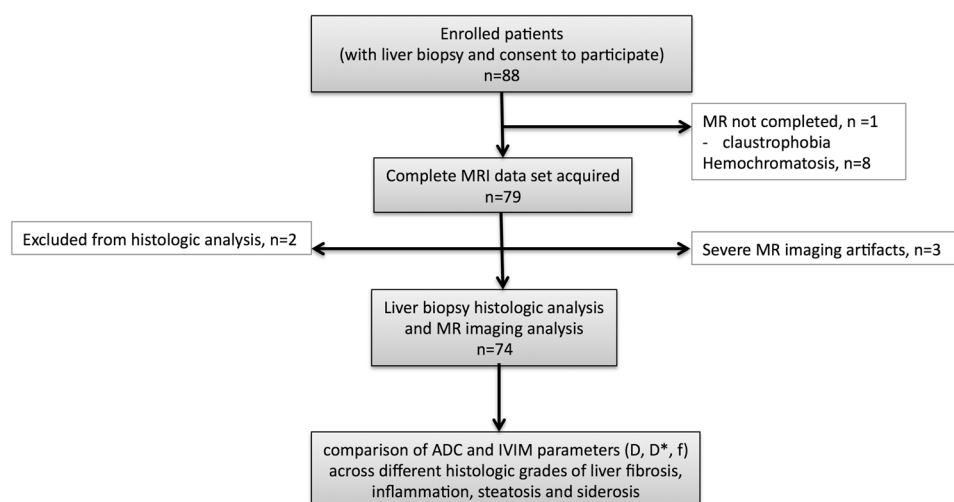


Fig. 1. Study protocol flow chart.

evaluated the liver samples using hematoxylin-eosin, Masson's trichrome, and Perls' Prussian stains. The Ishak modified scale was used for grading inflammation (0–18) and fibrosis (F0–F6) [28]. Hepatic steatosis was estimated according to the proportion of hepatocytes containing fat vesicles: no steatosis (grade 0, less than 5%); mild steatosis (grade 1, 5% to 33%); moderate steatosis (grade 2, 33% to 66%); and severe steatosis (grade 3, more than 66%) [29]. Iron loading was scored using a five grading system (0: no iron overload to 4: severe overload) [30].

MR imaging examination

Mean time interval between biopsy and MR examination was 2 days (range 0–25 days), most patients (77%) being examined in the same day of liver biopsy.

MR scans were performed in a 3 Tesla system (3T-TX Achieva, Philips Healthcare, The Netherlands), with a sixteen-channel phased-array coil. The liver MR imaging protocol included transverse T2-weighted turbo spin-echo sequence (TR/TE 850/80 ms, matrix 560, field of view $400 \times 285 \times 260$ mm, slice thickness 5 mm, number of slices 35), coronal transverse T2-weighted turbo spin-echo sequence (TR/TE 1170/80 ms, matrix 448, field of view $375 \times 448 \times 158$ mm, slice thickness 4 mm, number of slices 35), 2D multiecho chemical shift encoded gradient echo sequence (12 echoes, TE's = 0.99–8.69, echo spacing = 0.7 ms; TR = 10 ms; flip angle 10° ; voxel dimensions, 3×3 mm; slice thickness, 7 mm; 0.3 mm gap; field of view, 375×302 mm, single breath-hold acquisition) and transverse IVIM DW imaging sequence. IVIM imaging was performed using a single-shot spin-echo echo-planar sequence (SE-EPI), with gradient reversal fat suppression, using 6 b -values (0, 50, 200, 400, 600, and 800 s/mm^2) to maximize information in a reasonable time frame, and the following parameters: TR/TE 2205/74 ms, echo-planar imaging factor 63, FOV $375 \times 305 \times 207$ mm,

slice thickness 7 mm, matrix 256, receiver bandwidth 2000 Hz per pixel, number of slices 27, number of signals averaged 2, parallel imaging (accelerating factor 2) and respiratory-triggered acquisition. The scan duration was about 10 min.

Image analysis

Images were exported as raw data to calculate liver apparent diffusion coefficient (ADC) and IVIM parameters using QUIBIM research analysis platform (Valencia, Spain). Images from the different b -values were spatially registered using $b = 0$ as the reference. The ADC was calculated by a monoexponential fitting using all of b -values. The pure D and perfusion-related D^* and f parameters were obtained using the biexponential fitting [12].

One radiologist (MF, 8 years' experience) reviewed the images blinded to the histological results and clinical status. Three ROIs, 5 mm diameter, were manually placed in the right lobe and within the biopsied liver segment. ROIs were drawn in b -value = 0 and the coordinates of the voxels within the ROI were used to extract the values from the parametric maps. Care was taken to avoid blood vessels, focal lesions, motion artifacts, regions with poor signal to noise and biopsy tracts. For each ROI, the analysis was performed voxelwise and all the histogram was considered, the average of the three ROIs' values being used for statistical analysis as the dependent MR derived variables.

Reproducibility and repeatability tests

A test–retest experiment was designed to assess reproducibility and repeatability. Ten healthy volunteers (4 men, 6 women; mean age 33 years) were scanned with the same MR protocol as the patients. The MR examina-

tions were performed on two separate occasions, within different sessions at the same day. Three blinded observers (MF, radiologist with 8 years' experience; JO, radiologist in the fourth training year; JA, radiologist in the second training year) separately and independently evaluated the acquired images, placing 4 different ROIs in the right liver lobe.

In order to further assess intra-observer reproducibility, 10 randomly selected patients, with different grades of fibrosis and inflammatory activity, were assessed by re-reading the studies on a second separate occasion, by the same observer (MF).

Statistical analysis

Descriptive statistics were used to summarize the patients' characteristics. Categorical variables were evaluated by χ^2 and Fisher's exact test. For statistical analysis, necro-inflammatory activity score was grouped as none (score 0), mild (scores 1–6), moderate (scores 7–12), and severe inflammation (scores 13–18). The Ishak fibrosis staging was categorized into none (F0), mild to moderate fibrosis (F1–F3), and marked fibrosis/cirrhosis (F4–F6). Steatosis was grouped as none (grade 0), mild (grade 1) and significant steatosis (grade 2–3). Iron grades were categorized as none (grade 0), mild (grade 1), and moderated to severe siderosis (grade 2–4).

Differences in ADC, D , D^* , and f measurements across the categorical histologic grades of liver inflammation, fibrosis, steatosis or iron deposits were assessed with one-way analysis of variance, with post hoc Tukey HSD test. The Kruskal–Wallis test, with post hoc Mann–Whitney test and Bonferroni correction, was used when the variances were non-homogeneous, as was the case of f values across different grades of steatosis. Factorial analysis of variance was performed to evaluate the joint effect of fibrosis and inflammation. The Spearman correlation coefficient was used to calculate the degree of association between continuous and/or ordinal variables. Receiver operating characteristic (ROC) analysis evaluated the measurements' diagnostic performance. The area under the ROC curve (AUROC) was calculated for each set of dichotomized groups of histological grades of fibrosis (F0 vs. F1 or higher; F0–F3 vs. F4–F6) and necro-inflammatory activity (grade 0 vs. 1 or higher; grade 6 or lower vs. 7 or higher).

For measurements in volunteers, the agreement between the readers was assessed by intra-class correlation coefficients (ICCs) with 95% confidence intervals for the average values on each acquisition. Repeatability was evaluated between the first and second acquisitions in volunteers.

SPSS (version 22; SPSS, IBM, Chicago, Ill) was used for analysis. For all tests, a two-tailed p value of less than 0.05 was considered statistically significant.

Results

There were 74 patients in the final study population, including 38 men and 36 women, with a mean age of 46.4 years (standard deviation, 14.0; range of 19–77 years). The clinical indications for liver core biopsies were varied, including liver graft dysfunction ($n = 23$), persistent elevation of liver enzymes ($n = 17$), chronic HCV infection ($n = 11$), chronic alcoholic liver disease ($n = 5$), chronic HBV infection ($n = 5$), assessment of liver parenchyma during investigation of primary malignant focal liver lesion ($n = 3$), auto-immune hepatitis ($n = 3$), non-alcoholic fatty liver disease ($n = 3$), Wilson disease ($n = 3$), and toxic hepatitis ($n = 1$).

Histological results

All patients had adequate size biopsies (more than 15 mm in size). The stages of fibrosis and grades of necro-inflammatory activity, steatosis, and siderosis observed on histological analysis are summarized in Table 1.

Only 4 patients (5%) did not have liver fibrosis or necro-inflammatory activity. None of the patients had severe liver inflammation. Twenty-eight (37.8%) patients had neither liver steatosis nor iron overload, while 46 (62.2%) patients had both fat and iron liver deposition within the sample. The grades of steatosis and siderosis, the stage of fibrosis, and the score of necro-inflammatory activity were not significantly associated (Fisher's exact test, $p > 0.05$).

Liver ADC and IVIM measurements

The ADC, D , D^* , and f measurements are summarized in Table 2. These biomarkers are shown accordingly to the

Table 1. Histological characteristics of patient population ($n = 74$ patients)

Characteristic	Number (%)
Steatosis	
0	50 (67.6)
1	14 (18.9)
2	8 (10.8)
3	2 (2.7)
Iron	
0	39 (52.7)
1	21 (28.4)
2	11 (14.9)
3	2 (2.7)
4	1 (1.4)
ISHAK necro-inflammatory activity	
None (0)	6 (8.1)
Mild (1–6)	55 (74.3)
Moderate (7–12)	13 (17.6)
Severe (13–18)	0 (0)
ISHAK fibrosis grade	
None or minimal (F0–F1)	17 (23.0)
Moderate (F2–F3)	45 (60.8)
Marked fibrosis/cirrhosis (F4–F6)	12 (16.2)

Table 2. The ADC, D , D^* , and f measurements are summarized according to the histological hepatic fibrosis stage and necro-inflammatory activity, steatosis and siderosis grades

	ADC ($\times 10^{-3} \text{mm}^2/\text{s}$)			D ($\times 10^{-3} \text{mm}^2/\text{s}$)			D^* ($\times 10^{-3} \text{mm}^2/\text{s}$)			f (fraction)		
Study population average	1.62 (0.38)	1.52, 1.38–1.81		0.68 (0.22)	0.75, 0.56–0.82		67.99 (22.26)	70.13, 50.59–83.16		0.32 (0.08)	0.32, 0.26–0.37	
Fibrosis												
F0	1.83 (0.45)	1.71, 1.45–2.15		0.66 (0.19)	0.70, 0.55–0.80		62.16 (18.67)	68.24, 46.89–74.41		0.37 (0.08)	0.36, 0.33–0.40	
F1–F3	1.59 (0.34)	1.52, 1.38–1.79		0.68 (0.23)	0.76, 0.50–0.85		70.19 (23.46)	71.40, 51.92–86.38		0.31 (0.08)	0.31, 0.26–0.35	
F4–F6	1.41 (0.26)	1.37, 1.21–1.65		0.70 (0.21)	0.76, 0.60–0.82		67.96 (22.48)	71.04, 49.63–84.97		0.28 (0.05)	0.28, 0.22–0.33	
Necro-inflammatory activity												
0	1.94 (0.49)	1.85, 1.50–2.32		0.59 (0.14)	0.57, 0.45–0.73		75.62 (13.32)	75.36, 61.74–90.09		0.39 (0.09)	0.39, 0.31–0.46	
1–6	1.63 (0.38)	1.52, 1.40–1.84		0.67 (0.23)	0.76, 0.54–0.84		69.62 (21.19)	71.83, 51.20–85.86		0.32 (0.08)	0.33, 0.26–0.37	
7–12	1.43 (0.19)	1.46, 1.28–1.56		0.73 (0.18)	0.76, 0.61–0.86		57.54 (27.54)	57.45, 33.94–71.93		0.28 (0.05)	0.29, 0.25–0.32	
Steatosis												
0 (<5%)	1.60 (0.39)	1.49, 1.35–1.81		0.73 (0.19)	0.78, 0.64–0.86		67.24 (23.94)	70.81, 46.51–83.16		0.30 (0.08)	0.28, 0.25–0.33	
1 (6% to 33%)	1.78 (0.41)	1.74, 1.60–1.92		0.60 (0.22)	0.66, 0.43–0.80		70.13 (16.85)	70.40, 53.41–86.91		0.37 (0.07)	0.36, 0.32–0.43	
2 (>33%)	1.47 (0.41)	1.45, 1.33–1.52		0.47 (0.15)	0.45, 0.33–0.60		68.72 (22.79)	65.22, 48.32–81.32		0.35 (0.02)	0.34, 0.33–0.36	
Iron deposits												
0	1.63 (0.38)	1.53, 1.40–1.82		0.75 (0.18)	0.77, 0.63–0.86		63.73 (22.54)	68.03, 44.97–79.80		0.31 (0.08)	0.30, 0.26–0.34	
1	1.61 (0.32)	1.52, 1.42–1.70		0.69 (0.17)	0.76, 0.57–0.80		69.28 (18.24)	70.45, 52.57–83.78		0.32 (0.06)	0.33, 0.27–0.37	
2–4	1.58 (0.48)	1.46, 1.27–1.86		0.47 (0.25)	0.39, 0.27–0.76		77.91 (24.91)	76.34, 64.78–89.92		0.35 (0.10)	0.36, 0.31–0.41	

Group data are presented as mean (standard deviation), median and interquartile range

hepatic fibrosis stage and necro-inflammatory activity, and also to the steatosis and siderosis grades.

Liver fibrosis and inflammation

ADC and f values showed significant differences ($p = 0.009$, $p = 0.006$, respectively) across fibrosis stages (Fig. 2; Table 2): increasing stages of fibrosis showed lower ADC ($R_S = -0.32$, $p < 0.05$) and f values ($R_S = -0.38$, $p < 0.05$). Using Tukey HSD test, ADC and f measurements were only significantly different between patients with severe fibrosis or cirrhosis (F4–F6) and patients without fibrosis (F0) ($p < 0.05$).

Calculated ADC and f values were also different across histologic grades of liver inflammation ($p = 0.02$ and $p = 0.017$, respectively) (Fig. 3; Table 2). Using post hoc analysis, the ADC and f values were only significantly different between patients with moderate inflammation and patients without inflammation ($p < 0.05$), but were not significantly different between patients with none or mild inflammation ($p > 0.05$). ADC and f measurements had weak correlations with liver necro-inflammatory activity grading ($R_S = -0.31$, $p < 0.05$; $R_S = -0.32$, $p < 0.05$; respectively).

Neither the D nor the D^* values were significantly different across groups of liver fibrosis or inflammation ($p > 0.05$).

Since the ADC and f measurements were both affected by liver fibrosis and inflammation, we evaluated their joint effect with Factorial Analysis of Variance (dependent variable: ADC and IVIM parameters; factors: histological categorical variables—inflammation and fibrosis). The interaction effect between inflammation and fibrosis was not significant ($p = 0.748$); thus, it was removed from the univariate model. Considering the

main effects of both inflammation and fibrosis, fibrosis presented a significant effect on ADC ($p = 0.032$) and f measurements ($p = 0.031$), whereas inflammation had no significant effect neither on ADC ($p = 0.124$) nor f measurements ($p = 0.137$) (Fig. 4).

The accuracy of ADC or f measurements to discriminate the different stages of fibrosis or the necro-inflammatory activity grades was only fair. To distinguish between patients without fibrosis (F0) from patients with fibrosis (F1 or higher), the AUROC for ADC and f measurements were 0.670 and 0.740 ($p < 0.05$), respectively. To distinguish between patients without or mild fibrosis (F0–F3) from patients with severe fibrosis and cirrhosis (F4–F6), the AUROC for ADC and f values were 0.705 and 0.690 ($p < 0.05$), respectively. To differentiate patients without inflammation (score 0) from patients with necro-inflammatory activity (score 1 or higher), the AUROC for ADC and f values were 0.748 and 0.749 ($p < 0.05$), respectively. To distinguish patients with none or minor (scores 0–6) from moderate inflammatory activity (scores 7–12), the AUROC for ADC and f values were 0.682 and 0.687 ($p < 0.05$), respectively.

Hepatic steatosis and siderosis

Calculated D measurements were significantly different across the histologic grades of steatosis ($p < 0.001$) and iron overload ($p < 0.001$), whereas f measurements showed significant differences only across histologic steatosis grades (Kruskal–Wallis, $p = 0.005$) (Fig. 5). Patients with increased fat deposition showed lower D values ($R_S = -0.44$, $p < 0.001$) and higher f values ($R_S = 0.42$, $p < 0.001$). However, the mean f value of grade 1 steatosis group (0.37) was higher than grade 0 group (0.30) and grade 2 group (0.35) (Table 2).

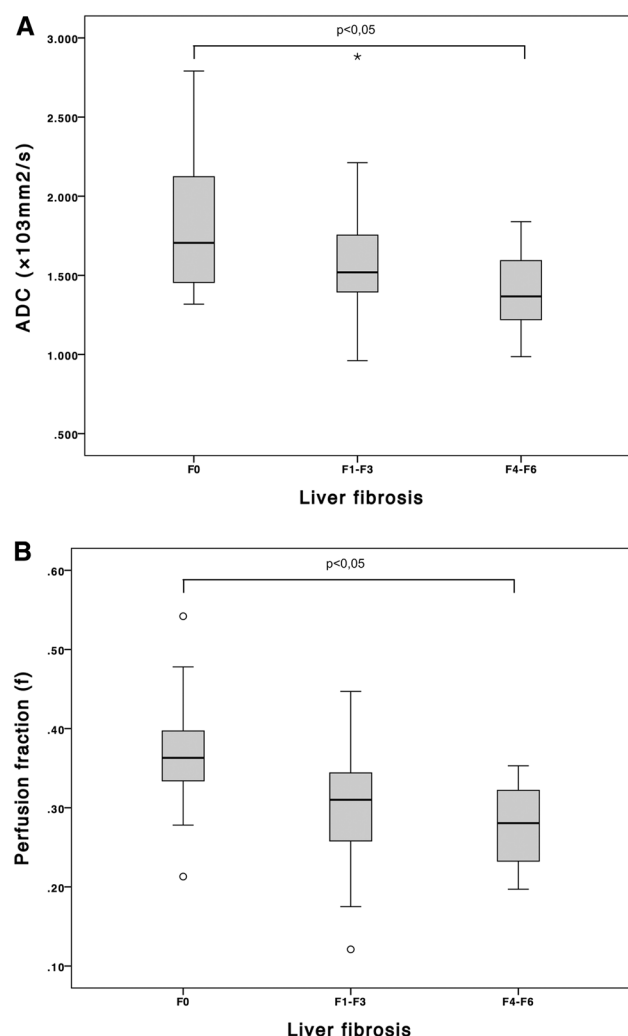


Fig. 2. Box and whisker plots of ADC (**A**) and perfusion fraction (f) (**B**) measured in all patients, compared with histological fibrosis staging. **A** There was a negative correlation in ADC values with increasing Ishak fibrosis stage ($R_S = -0.32$, $p < 0.05$). **B** There was a negative correlation in perfusion fraction (f) with increasing Ishak fibrosis stage ($R_S = -0.38$, $p < 0.05$).

The D values were negatively correlated with iron overload ($R_S = -0.36$, $p = 0.002$).

There were no significant differences on the ADC and D^* values among histologic grades of liver steatosis or siderosis ($p > 0.05$).

Unfortunately, the number of patients in this study did not allow a multivariate analysis to test for all the interactions between the different categorical histologic variables in ADC or IVIM measurements.

Reproducibility of DW measurements

The median, IQR for ADC and IVIM values of healthy volunteers are presented in Table 3. Agreement was considered as poor (ICC, 0–0.50), fair to good (ICC,

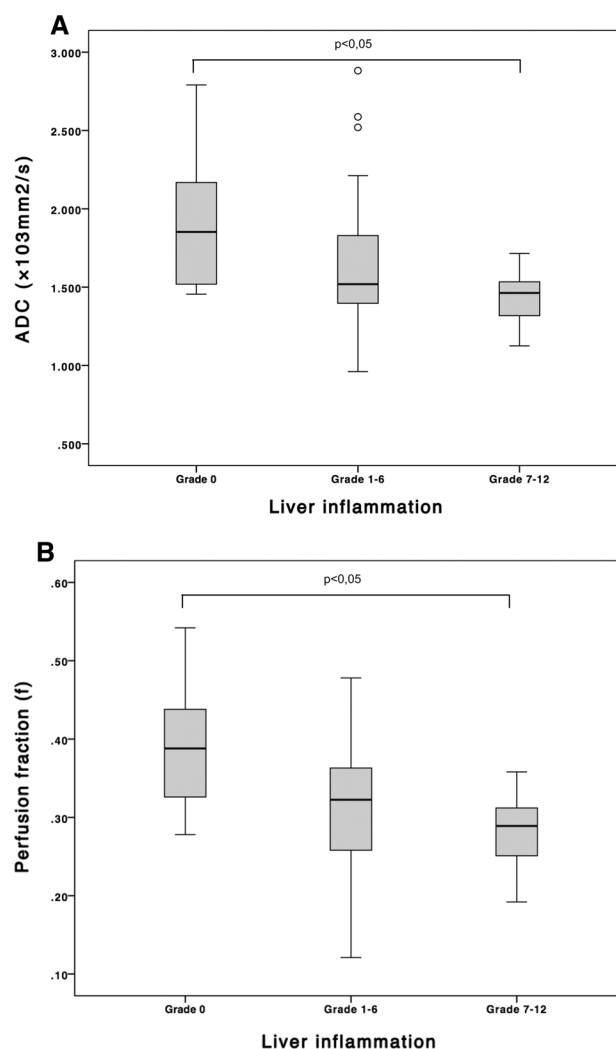


Fig. 3. Box and whisker plots of ADC (**A**) and perfusion fraction (f) values (**B**) measured in all patients, compared with histologic necro-inflammatory activity grading. **A** There was a negative correlation in ADC values with increasing Ishak necro-inflammatory grading ($R_S = -0.31$, $p < 0.05$). **B** There was a negative correlation in perfusion fraction with increasing Ishak necro-inflammatory grading ($R_S = -0.32$, $p < 0.05$).

0.50–0.80), very good (ICC > 0.80) or excellent (ICC, > 0.90).

Regarding the test–retest examination of 10 volunteers, there was an excellent agreement between the different readers for the measurements of ADC, D and f , and a fair agreement for the D^* measurements. The reproducibility between the two different MR examinations was good for the ADC measurements (ICC > 0.8, $p < 0.05$) and excellent for the D and f estimated values (ICC > 0.86, $p < 0.05$). However, for the D^* measurements, there was no agreement between measurements from the two different examinations ($p > 0.05$).

The intra-reader variability in the 10 patients was assessed by the ICCs between measurements performed

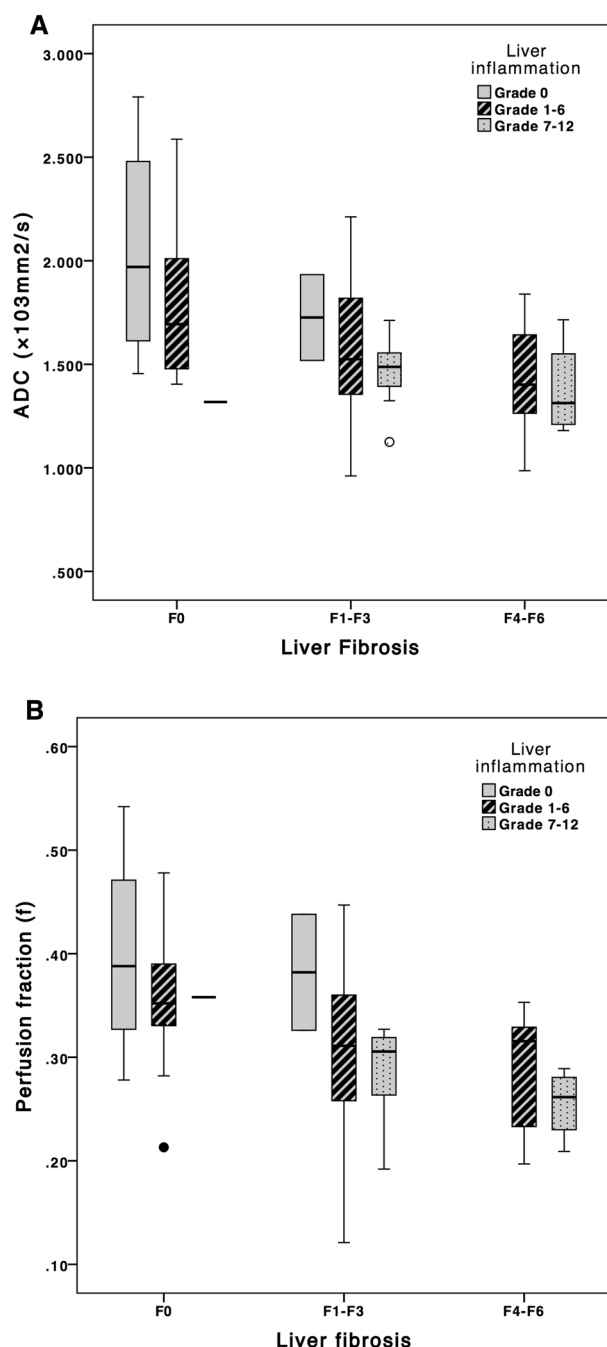


Fig. 4. Box and whisker plots of ADC (**A**) and perfusion fraction (f) values (**B**) measured in all patients, compared with histologic fibrosis staging and grouped by necro-inflammatory activity grading. Considering the main effects of both inflammation and fibrosis, fibrosis presented a significant effect on ADC ($p = 0.032$) and f measurements ($p = 0.031$), with lower ADC in higher stages of liver fibrosis, whereas inflammation had no significant effect neither on ADC ($p = 0.124$) nor f measurements ($p = 0.137$).

on two different occasions. The intra-reader ICC on volunteers was excellent ($ICC > 0.91$, $p < 0.05$) for the ADC and all IVIM measurements. Furthermore, the

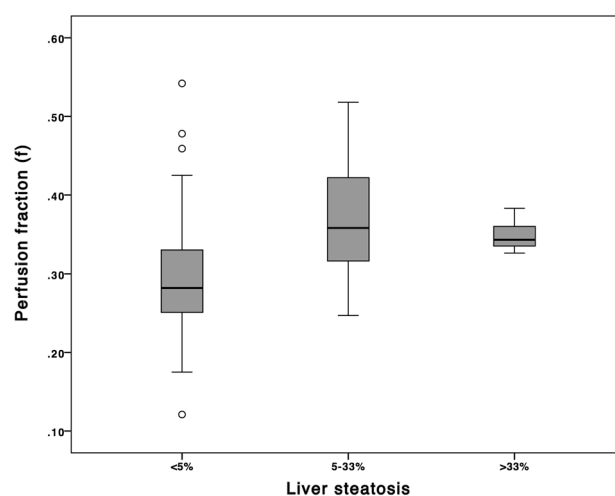


Fig. 5. Box and whisker plots of perfusion fraction (f) values measured in all patients, compared with histologic grading of liver steatosis (percentage of hepatocytes with cytoplasmatic lipid vacuoles). Higher perfusion fraction was observed in patients with increased fat deposition ($R_s = 0.42$, $p < 0.001$).

intra-reader ICC in 10 patients was excellent for all of the MR measurements ($ICC > 0.92$, $p < 0.001$).

Discussion

We report that ADC and f values were lower with higher fibrosis stages, in patients with diffuse liver diseases. However, the correlation between these MR measurements and fibrosis stages was weak and the accuracy of ADC or f measurements to discriminate the different stages of fibrosis or the necro-inflammatory activity grades was only fair.

Several studies have associated reduced liver ADC measurements with liver fibrosis and cirrhosis [4, 6–9, 13, 17, 20–22, 31, 32]. ADC measurements contain both diffusion and perfusion components, which difficult the biological interpretation of reduced ADC values. The IVIM model allows separating the liver signal intensity changes due to true molecular diffusion from those related to perfusion within the tissue. Recently, it has been proposed that the IVIM model, and particularly the perfusion-related parameters D^* and f , would be more advantageous to evaluate liver fibrosis than conventional ADC measurements [5, 13–18, 33–35]. However, the influence of liver inflammation, steatosis and siderosis on liver values has to be considered in this analysis.

Initial studies showed that D^* values measurements were lower in patients with cirrhosis than in non-cirrhotic patients [5]. However, other studies observed lower D , D^* and f values in patients with cirrhosis compared to non-cirrhotic patients [33], and lower f and D^* values in patients with severe fibrosis (METAVIR $\geq F3$) [17]. Lower D , D^* , and f values were observed with increasing stages of liver fibrosis in patients studied at 1.5

Table 3. MR measurements data on 10 volunteers, with the analysis of inter-examination, inter-reader and intra-reader variability

	ADC ($\times 10^{-3}$ mm ² /s)	<i>D</i> ($\times 10^{-3}$ mm ² /s)	<i>D</i> * ($\times 10^{-3}$ mm ² /s)	<i>f</i>
Average values in volunteers ^a	1.82, 1.53–2.27	0.73, 0.47–0.86	69.99, 63.41–77.16	0.36, 0.31–0.45
Inter-examination variability ^b				
Observer 1	0.83 (0.32–0.96)	0.96 (0.86–0.99)	−0.06 (−5.53–0.76)	0.93 (0.75–0.98)
Observer 2	0.87 (0.51–0.97)	0.92 (0.68–0.98)	−1.38 (−7.83–0.42)	0.86 (0.48–0.96)
Observer 3	0.91 (0.64–0.98)	0.96 (0.86–0.99)	0.55 (−0.66–0.89)	0.94 (0.70–0.98)
Inter-observer variability ^b				
Time 1	0.95 (0.86–0.99)	0.99 (0.97–1.00)	0.77 (0.36–0.94)	0.98 (0.93–0.99)
Time 2	0.93 (0.80–0.98)	0.97 (0.93–0.99)	0.70 (0.21–0.92)	0.95 (0.86–0.99)
Intra-reader variability (10 patients) ^b	0.94 (0.63–0.99)	0.98 (0.93–1.00)	0.93 (0.75–0.98)	0.92 (0.70–0.98)

^a Median, IQR^b ICC (95% confidence interval)

and 3 Tesla magnets [13, 14]. Three different animal studies reported contradictory results. Chow et al. [16] observed lower *D* and *D** values with increasing liver fibrosis. Zhang et al. [18] observed that *f* values decreased with increasing fibrosis stage; however, they did not assess *D** parameter. Finally, Hu et al. [15] observed inverse correlations between fibrosis stages and *D*, *D**, and *f* measurements. Unluckily, none of these animal studies evaluated the influence of necro-inflammatory activity, fat or iron on IVIM parameters.

Our findings of diminished ADC and *f* values with increasing fibrosis support the previously reported perfusion-related association. As ADC measurements contain both diffusion and perfusion components, the different results for ADC and *D* values might be explained by changes in the liver perfusion rather than pure diffusion of water molecules restriction. These findings most probably reflect the accumulation of collagen deposits, which increases the hepatic resistance to portal blood flow and reduces the portal perfusion within the liver.

Inflammatory changes, such as cell necrosis and inflammatory cell infiltration within the sinusoids, might also influence DW signal. It has been reported that ADC and *f* values decreased not only with fibrosis stage [4–9, 13, 17, 20–22, 31–33] but also with the severity of necro-inflammatory activity [6, 7, 9, 20–22]. However, in our series, only fibrosis was the main explicative variable for the ADC and perfusion-related *f* decrements.

Although we found differences across grades of inflammation and fibrosis with ADC and *f* measurements, only a mild correlation was discovered between the histological grades and MR measurements. Moreover, the AUROC that evaluate dichotomized fibrosis stages or inflammatory grades were fair, under 0.8. Thus, neither the ADC nor the *f* measurements demonstrated adequate accuracy to discriminate between different stages of liver fibrosis or grades of inflammatory activity. Therefore, the observed differences do not allow for an appropriate staging of fibrosis and inflammation pathological markers.

Imaging biomarkers might also be influenced by confounding factors such as fat or iron deposits, which

are frequently present in diffuse liver diseases. We observed that pure *D* decreased with increasing grades of steatosis, similar with previous reports [24, 27, 34, 35]. Steatosis decrease water molecules movement due to the enlarged hepatocytes, overloaded with macrovesicles of fat, restricting both intracellular and extracellular diffusion [24]. Furthermore, the *f* measurements were influenced by fat deposition, increasing with the grade of liver steatosis. Steatosis may have a confounding effect in *f* measurements. The mean *f* values of grade 1 steatosis group (0.37) were higher than grade 0 (0.30) and grade 2 (0.35) groups, which might be explained by the effect of fibrosis, inflammation and steatosis in *f* measurements. Unfortunately, as previously mentioned, the number of patients in this study did not allow a multivariate analysis to test for all the interactions between the different categorical histologic variables in ADC or IVIM measurements.

Only the *D* values were significantly reduced by iron deposits, decreasing with the severity of overload. These values decreased mainly due to the signal decay and noise floor effects induced by the susceptibility effects of iron overload [23, 36].

Regarding *D** measurements, they were not related to the stage of liver fibrosis or inflammation and, quite important, they were not consistent between different MR examinations in volunteers.

This study has some limitations. We used 6 different *b*-values, a lower number than some previous studies. Nevertheless, there is no consensus about the ideal number of *b*-values that should be used and, moreover, it has been demonstrated the feasibility of IVIM analysis of DW images using a number of *b*-value as low as 4 [17, 37]. We have assessed test–retest reproducibility in volunteers and not in patients with liver diseases; the obtained results from normal tissue might not be adequately extrapolated to diseased livers [38]. There were relatively fewer patients with severe steatosis or siderosis scores, and no patient with severe inflammation, which might have limited the statistical significance of the results and have hampered the performance of factorial analyses with all the variables together. Another limiting factor is that the MR images and the histological

results were analyzed by only one reader and, therefore, variability of both these results could not be assessed. ROIs were drawn in the same liver segment than liver biopsy to minimize the effect of the heterogeneously distributed diffuse liver damage. As most MR examinations were performed after liver biopsy, care was taken to avoid the biopsy tract although unrecognized hemorrhage or edema might have influenced MR results. Also, Ishak score is not a continuous scale and the increase of fibrous tissue accumulation or necro-inflammatory activity from one grade to another is not linear [39]. However, even with this limitations, liver biopsy is still considered the best reference standard. Although larger scale studies should validate the relationship between ADC and IVIM parameters and different degrees of fibrosis and inflammation, to definitively evaluate the interactions of different histologic variables on IVIM parameters, we believe that after our study the role of IVIM is limited in routine clinical practice and patient's outcome evaluation.

In conclusion, in spite of the differences in ADC and f derived IVIM values across different groups of liver fibrosis and inflammation, these metrics cannot properly discriminate between different stages of fibrosis or inflammation in different diffuse liver diseases. Steatosis may confound the f measurements. D^* and D parameters are not useful biomarkers of liver fibrosis: D^* is not a reproducible measure and D measurements are influenced by confounders such as fat and iron, which commonly coexist with liver fibrosis.

Acknowledgments. This work was partially funded by a research Grant from the Teaching and Research Department of Centro Hospitalar do Porto (DEFI:309/12(213-DEFI/251-CES)) and from a Spanish Ministry of Health and Carlos III Health Institute funding grant (PI12/01262). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Compliance with ethical standards

Conflicts of interest Javier Sanchez Gonzalez is employee at Philips Healthcare Iberia. Angel Alberich Bayarri and Luis Marti-Bonmati are co-founders of QUIBIM SME. The remaining authors declare that they have no conflicts of interest.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent Written informed consent was obtained from all patients.

References

- Ellis EL, Mann DA (2012) Clinical evidence for the regression of liver fibrosis. *J Hepatol* 56:1171–1180. doi:10.1016/j.jhep.2011.09.024
- Myers RP, Fong A, Shaheen AAM (2008) Utilization rates, complications and costs of percutaneous liver biopsy: a population-based study including 4275 biopsies. *Liver Int* 28:705–712. doi:10.1111/j.1478-3231.2008.01691.x
- Bedossa P, Dargère D, Paradis V (2003) Sampling variability of liver fibrosis in chronic hepatitis C. *Hepatology* 38:1449–1457. doi:10.1016/j.jhep.2003.09.022
- Lewin MM, Poujol-Robert AA, Boëlle P-YP, et al. (2007) Diffusion-weighted magnetic resonance imaging for the assessment of fibrosis in chronic hepatitis C. *Hepatology* 46:658–665. doi:10.1002/hep.21747
- Luciani A, Vignaud A, Cavet M, et al. (2008) Liver cirrhosis: Intravoxel incoherent motion MR Imaging—Pilot Study. *Radiology* 249:891–899. doi:10.1148/radiol.2493080080
- Taouli B, Tolia AJ, Losada M, et al. (2007) Diffusion-weighted MRI for quantification of liver fibrosis: preliminary experience. *AJR Am J Roentgenol* 189:799–806. doi:10.2214/AJR.07.2086
- Taouli B, Chouli M, Martin AJ, et al. (2008) Chronic hepatitis: Role of diffusion-weighted imaging and diffusion tensor imaging for the diagnosis of liver fibrosis and inflammation. *J Magn Reson Imaging* 28:89–95. doi:10.1002/jmri.21227
- Sandrasegaran K, Akisik FM, Lin C, et al. (2009) Value of diffusion-weighted MRI for assessing liver fibrosis and cirrhosis. *AJR Am J Roentgenol* 193:1556–1560. doi:10.2214/AJR.09.2436
- Bakan AA, Inci E, Bakan S, et al. (2011) Utility of diffusion-weighted imaging in the evaluation of liver fibrosis. *Eur Radiol* 22:682–687. doi:10.1007/s00330-011-2295-z
- Taouli B, Koh DM (2009) Diffusion-weighted MR Imaging of the Liver. *Radiology* 254:47–66. doi:10.1148/radiol.09090021
- Yamada I, Aung W, Himeno Y, et al. (1999) Diffusion coefficients in abdominal organs and hepatic lesions: evaluation with intravoxel incoherent motion echo-planar MR imaging. *Radiology* 210:617–623
- Le Bihan D, Breton E, Lallemand D, et al. (1988) Separation of diffusion and perfusion in intravoxel incoherent motion MR imaging. *Radiology* 168:497–505. doi:10.1148/radiology.168.2.3393671
- Yoon JH, Lee JM, Baek JH, et al. (2014) Evaluation of hepatic fibrosis using intravoxel incoherent motion in diffusion-weighted liver MRI. *J Comput Assist Tomogr* 38:110–116. doi:10.1097/RCT.0b013e3182a589be
- Lu P-X, Huang H, Yuan J, et al. (2014) Decreases in molecular diffusion, perfusion fraction and perfusion-related diffusion in fibrotic livers: A prospective clinical intravoxel incoherent motion MR imaging study. *PLoS One* 9:e113846. doi:10.1371/journal.pone.0113846.t002
- Hu G, Chan Q, Quan X, et al. (2014) Intravoxel incoherent motion MRI evaluation for the staging of liver fibrosis in a rat model. *J Magn Reson Imaging* 42:331–339. doi:10.1002/jmri.24796
- Chow AM, Gao DS, Fan SJ, et al. (2012) Liver fibrosis: an intravoxel incoherent motion (IVIM) study. *J Magn Reson Imaging* 36:159–167. doi:10.1002/jmri.23607
- Chung SR, Lee SS, Kim N, et al. (2014) Intravoxel incoherent motion MRI for liver fibrosis assessment: a pilot study. *Acta Radiol* 56:1428–1436. doi:10.1177/0284185114559763
- Zhang Y, Jin N, Deng J, et al. (2013) Intra-voxel incoherent motion MRI in rodent model of diethylnitrosamine-induced liver fibrosis. *Magn Reson Imaging* 31:1017–1021. doi:10.1016/j.mri.2013.03.007
- Saxena R (2011) Microscopic anatomy, basic terms, and elemental lesions. In: Saxena R (ed) *Practical hepatic pathology: a diagnostic approach*, 1st edn. Philadelphia: Elsevier Saunders, pp 3–28
- Wang Q-B, Zhu H, Liu H-L, Zhang B (2012) Performance of magnetic resonance elastography and diffusion-weighted imaging for the staging of hepatic fibrosis: a meta-analysis. *Hepatology* 56:239–247. doi:10.1002/hep.25610
- Tosun M, Inan N, Sarisoy HT, et al. (2013) Diagnostic performance of conventional diffusion weighted imaging and diffusion tensor imaging for the liver fibrosis and inflammation. *Eur J Radiol* 82:203–207. doi:10.1016/j.ejrad.2012.09.009
- Fujimoto K, Tonan T, Azuma S, et al. (2011) Evaluation of the mean and entropy of apparent diffusion coefficient values in chronic hepatitis C: correlation with pathologic fibrosis stage and inflammatory activity grade. *Radiology* 258:739–748. doi:10.1148/radiol.10100853
- Bülow R, Mensel B, Meffert P, et al. (2012) Diffusion-weighted magnetic resonance imaging for staging liver fibrosis is less reliable

- in the presence of fat and iron. *Eur Radiol* 23:1281–1287. doi:[10.1007/s00330-012-2700-2](https://doi.org/10.1007/s00330-012-2700-2)
24. Guiu B, Petit JM, Capitan V, et al. (2012) Intravoxel incoherent motion diffusion-weighted imaging in nonalcoholic fatty liver disease: a 3.0-T MR Study. *Radiology* 265:96–103. doi:[10.1148/radiol.12112478](https://doi.org/10.1148/radiol.12112478)
 25. Lee JT, Liao J, Murphy P, et al. (2012) Cross-sectional investigation of correlation between hepatic steatosis and IVIM perfusion on MR imaging. *Magn Reson Imaging* 30:572–578. doi:[10.1016/j.mri.2011.12.013](https://doi.org/10.1016/j.mri.2011.12.013)
 26. Hansmann J, Hernando D, Reeder SB (2012) Fat confounds the observed apparent diffusion coefficient in patients with hepatic steatosis. *Magn Reson Med* 69:545–552. doi:[10.1002/mrm.24535](https://doi.org/10.1002/mrm.24535)
 27. Leitão HS, Doblas S, d'Assignies G, et al. (2012) Fat deposition decreases diffusion parameters at MRI: a study in phantoms and patients with liver steatosis. *Eur Radiol* 23:461–467. doi:[10.1007/s00330-012-2626-8](https://doi.org/10.1007/s00330-012-2626-8)
 28. Ishak K, Baptista A, Bianchi L, et al. (1995) Histological grading and staging of chronic hepatitis. *J Hepatol* 22:696–699
 29. Kleiner DE, Brunt EM, Van Natta M, et al. (2005) Design and validation of a histological scoring system for nonalcoholic fatty liver disease. *Hepatology* 41:1313–1321. doi:[10.1002/hep.20701](https://doi.org/10.1002/hep.20701)
 30. Deugnier Y, Turlin B (2007) Pathology of hepatic iron overload. *WJG* 13:4755–4760
 31. Papalavrentios L, Sinakos E, Chourmouzi D, et al. (2015) Value of 3 Tesla diffusion-weighted magnetic resonance imaging for assessing liver fibrosis. *Ann Gastroenterol* 28:118–123
 32. Bonekamp S, Torbenson MS, Kamel IR (2011) Diffusion-weighted magnetic resonance imaging for the staging of liver fibrosis. *J Clin Gastroenterol* 45:885–892. doi:[10.1097/MCG.0b013e318223bd2c](https://doi.org/10.1097/MCG.0b013e318223bd2c)
 33. Patel J, Sigmund EE, Rusinek H, et al. (2010) Diagnosis of cirrhosis with intravoxel incoherent motion diffusion MRI and dynamic contrast-enhanced MRI alone and in combination: Preliminary experience. *J Magn Reson Imaging* 31:589–600. doi:[10.1002/jmri.22081](https://doi.org/10.1002/jmri.22081)
 34. Murphy P, Hooker J, Ang B, et al. (2014) Associations between histologic features of nonalcoholic fatty liver disease (NAFLD) and quantitative diffusion-weighted MRI measurements in adults. *J Magn Reson Imaging* 41:1629–1638. doi:[10.1002/jmri.24755](https://doi.org/10.1002/jmri.24755)
 35. Parente DB, Paiva FF, Oliveira Neto JA, et al. (2015) Intravoxel incoherent motion diffusion weighted MR imaging at 3.0 T: assessment of steatohepatitis and fibrosis compared with liver biopsy in type 2 diabetic patients. *PLoS One* 10:e0125653. doi:[10.1371/journal.pone.0125653.t004](https://doi.org/10.1371/journal.pone.0125653.t004)
 36. Chandarana H, Do RKG, Mussi TC, et al. (2012) The effect of liver iron deposition on hepatic apparent diffusion coefficient values in cirrhosis. *AJR Am J Roentgenol* 199:803–808. doi:[10.2214/AJR.11.7541](https://doi.org/10.2214/AJR.11.7541)
 37. Leporq B, Saint-Jalmes H, Rabrait C, et al. (2015) Optimization of intra-voxel incoherent motion imaging at 3.0 Tesla for fast liver examination. *J Magn Reson Imaging* 41:1209–1217. doi:[10.1002/jmri.24693](https://doi.org/10.1002/jmri.24693)
 38. Andreou A, Koh DM, Collins DJ, et al. (2012) Measurement reproducibility of perfusion fraction and pseudodiffusion coefficient derived by intravoxel incoherent motion diffusion-weighted MR imaging in normal liver and metastases. *Eur Radiol* 23:428–434. doi:[10.1007/s00330-012-2604-1](https://doi.org/10.1007/s00330-012-2604-1)
 39. Desmet VJ (2003) Knodell RG, Ishak KG, Black WC, Chen TS, Craig R, Kaplowitz N, Kiernan TW, Wollman J. Formulation and application of a numerical scoring system for assessing histological activity in asymptomatic chronic active hepatitis [Hepatology 1981;1:431–435]. *J Hepatol* 38:382–386. doi:[10.1016/S0168-8278\(03\)00005-9](https://doi.org/10.1016/S0168-8278(03)00005-9)